

ESTUDIOS

EL USO DE LA ETIQUETA META EN SITIOS WEB DE BIBLIOTECAS

José A. Senso y Luis Villén Rueda*

Resumen: Cada vez es más común la utilización de técnicas de posicionamiento para lograr una mayor visibilidad de sitios web. Una de las herramientas para lograr este objetivo la proporciona el propio lenguaje HTML: la etiqueta META. En este trabajo se profundiza en su desarrollo y uso y se evalúa su uso dentro de servidores web de bibliotecas representativas en el ámbito nacional e internacional.

Palabras clave: metadatos, recuperación de información en el web, visibilidad de páginas corporativas, posicionamiento en motores de búsqueda, etiqueta meta.

Abstract: Positioning techniques are being used increasingly to obtain greater visibility for web sites. One of the tools for reaching this objective is provided by HTML itself: the META tags. This work details how they are developed and implemented, and includes an evaluation of their use by web servers from a sampling of national and international libraries.

Keywords: metadata, information retrieval on World Wide Web, visibility of corporate web pages, search engine positioning, meta tag.

1 Introducción

En la actualidad parece evidente lo vital que es para cualquier organización importante que su presencia en Internet no sólo sea conocida, sino que además se pueda localizar rápidamente y que toda la información que cuelga en sus servidores esté recuperable desde cualquier servicio de consulta de la Red.

Fundamentalmente existen dos factores a tener en cuenta para lograr que un sitio web sea conocido. Por un lado se encuentra la visibilidad, es decir, la facilidad con la que los clientes/usuarios/visitantes potenciales pueden encontrar un sitio web establecido de entre el resto de servidores de características similares. El otro factor, el posicionamiento, utiliza diversas técnicas para lograr que determinados webs se distribuyan en la parte más alta de la lista de resultados tras una consulta a un servicio de búsqueda. Ambos elementos, unidos a las necesidades empresariales de estar bien «ubicado» en Internet, han propiciado el creciente aumento de empresas que se encargan de algo tan simple —y a la vez tan complejo— como es el asegurar que un web será localizado siempre que se necesite.

Son numerosos los artículos en los que se defiende la idea de que la correcta introducción de elementos descriptivos en las etiquetas META del HTML facilita el trabajo de los servicios de búsqueda (especialmente los motores, como es obvio) y proporcionan un mejor posicionamiento (1, 2, 3 y 4), aunque no es menos cierto que

* Universidad de Granada. Facultad de Biblioteconomía y Documentación. Correo-e: jsenso@ugr.es; lvillen@ugr.es.

Recibido: 7-4-03; 2.^a versión: 30-3-04.

también existen otros autores que defienden la idea de que el uso de metaetiquetas no garantiza un posicionamiento óptimo (5 y 6).

A pesar de todo, es indudable que su presencia redundante en una mayor visibilidad del servidor, no sólo en lo que a información se refiere, sino también a la presencia corporativa dentro de la Red.

Así, en experimentos desarrollados por Turner y Brackbill (7) se demuestra que aumenta considerablemente la posibilidad de localizar sitios web de compañías que utilizan palabras clave en las metaetiquetas (en este estudio se habla sobre todo de *infoseek* —ahora desaparecido— y *Altavista*) con respecto a otras empresas que no hacen uso de estos elementos.

1.1 Uso de la etiqueta META

La etiqueta META se utiliza dentro de la cabecera (entre las etiquetas <HEAD> y </HEAD>) de una página HTML desde la versión 2.0 de este lenguaje (8) con la finalidad de *descubrir que los datos existen y cómo se puede acceder a ellos y para documentar el contenido, calidad y características de los datos, indicando su conveniencia de uso.*

Por la forma en la que operan los servicios de búsqueda de Internet, está claro que el primer objetivo propuesto para la creación de esta etiqueta no se podrá alcanzar en mucho tiempo. Pero recordemos: el espíritu que interesa para este trabajo es el que se desprende de la segunda parte de la «declaración de intenciones». El hecho de contemplar la posibilidad de establecer unos mecanismos que permitan identificar, indizar y catalogar recursos electrónicos permitió, por un lado, enseñar el camino a gran cantidad de sistemas de metadatos creados con posterioridad (Dublin Core, RDF —Resource Description Framework— y MCF —Meta Content Framework— especialmente) y que se beneficiaron de los avances de la introducción del trabajo con DTDs —Document Type Definitions—. Por otro lado, propiciaron la aparición de estudios que avalaban el uso de esta serie de mecanismos para facilitar la localización de información en un entorno tan cambiante como son todos aquellos que se basan en la arquitectura cliente/servidor (9).

De forma resumida, la manera de utilizar esta etiqueta es la siguiente:

HTTP-EQUIV. Especifica el encabezamiento de respuesta HTTP más pertinente para el tipo de información que contiene. Si el elemento utilizado para describir esta parte de la etiqueta es identificado por el servidor HTTP, el contenido se podrá procesar sobre la base de un sistema sintáctico reconocido. En caso de no encontrar este atributo, el servidor utilizará el atributo NAME para identificar la información codificada en la etiqueta META.

NAME. Se emplea para especificar el tipo de metadato utilizado para describir la información que contiene la página. Si este atributo no se encuentra presente, HTTP-EQUIV hará la misma función que él. Las propiedades que puede incluir son:

- *Author:* nombre de la persona con la responsabilidad intelectual del contenido del documento.
 <META NAME="Author" CONTENT="Scott Adams">

- *Description*: breve resumen del contenido del documento.
<META NAME="Description" CONTENT="Resumen de la página">
- *Generator*: especifica el nombre y la versión del programa de edición con el que se ha realizado la página. Se suele utilizar para que las empresas de software analicen el impacto de su programa.
<META NAME="Generator" CONTENT="Microsoft FrontPage 3.0">
- *Keywords*: lista de palabras clave que será utilizada por algunos motores de búsqueda para indizar el documento (además de aquellas que se extraigan automáticamente del cuerpo del mismo).
<META NAME="Keywords" CONTENT="palabra clave1, palabra clave2">
- *Robots*: controla la acción de los robots cuando se detengan en el documento. Las cuatro acciones que se puede especificar son: *index* (para indizar esta página, mientras *noindex* no diga lo contrario), *follow* (el robot seguirá todos los vínculos del documento —internos y externos— para indizarlos a no ser que *nofollow* especifique lo contrario), *all* (que equivale a *index*, *follow*, que es la opción utilizada por defecto) y *none* (*noindex* y *nofollow* a la vez) (1):
<META NAME="Robots" CONTENT="noindex, follow">

El atributo NAME permite la utilización de otras etiquetas como Classification, Copyright, Rating, Formatter, preferred browser, number of pages..., pero ninguna de ellas ha sido implementada con éxito por ningún navegador.

Content. Se trata de una ampliación del atributo HTTP-EQUIV que especifica el tipo de contenido al que se refiere éste o el atributo NAME. Esta información puede ser de muchos tipos, tal y como se puede observar en los ejemplos siguientes:

- Versión con la que se trabaja.
- URL para ampliar información.
- Correo electrónico del autor de la página.
- Realizar otra conexión.
- Actualizar una página cada X segundos.

El correcto uso de esta etiqueta tiene sentido con vistas a facilitar la recuperación por medio de los sistemas de búsqueda de Internet. Tras diversas consultas realizadas en foros virtuales (listas de distribución especialmente), nos hemos percatado de que uno de los principales escollos que se encuentra el uso de las etiquetas META del HTML deriva de la creencia de que son pocos los servicios de búsqueda en Internet que son capaces de hacer «algo» con ellas. Esta opinión difiere mucho de la realidad. Así, son varios los servicios de búsqueda que se valen, o se han valido, de los metadatos para realizar su trabajo (independientemente de si su uso es para Internet o para una Intranet). De hecho, no son únicamente los programas que trabajan con robots o arañas los que los utilizan. Algunos de ellos los encontramos en la siguiente lista que recoge tanto motores como directorios que existen o han existido en Internet y que, en algún momento de su «vida útil», más allá de adquisiciones, compras o quiebras, emplearon estas etiquetas:

- Yahoo (10): antes de emplear la base de datos Google utilizaba todas las etiquetas META de HTML para realizar la primera búsqueda general, en las siguientes consultas —lanzadas dentro del conjunto de resultados iniciales— ya no las apro-

vecha. Es incapaz de reconocer el atributo «refresh», por lo que si no se ha especificado previamente un valor «follow» interrumpirá la indización y no saltará a la URL indicada en su momento. Este comportamiento es común en otros servicios de Internet como AOL Search (10), Excite, Google, Hotbot (incluye los primeros 75 caracteres de META «keywords», y 150 de «description»), iWon (250 caracteres de «description») (10), Lycos (procesa las 20-15 primeras palabras aparecidas en «description»), MSN (las primeras 250 palabras en «description»), NetScape Search, Northern Light (200 caracteres de «description») (10) o Raging Search (1.017 caracteres en «keywords» y otros tantos en «description») (10).

- Go: hoy ya desaparecido, fue uno de los primeros buscadores considerados por los expertos como híbrido –usa un motor de búsqueda basado en una araña que se complementa con una estructura jerárquica de conceptos propia– que, además, indizaba los primeros 75 caracteres incluidos en la etiqueta <TITLE>, los 1.024 primeros aparecidos en META «keywords» y los primeros 246 símbolos incluidos en META NAME («description») (10).
- Altavista: al mostrar el resultado de una consulta, el buscador ofrece un resumen del contenido de cada documento. Si éste incluye la etiqueta META NAME="description/subject/abstract", el resumen recogerá el valor asociado al atributo CONTENT. Si el documento no contiene esta etiqueta, el resumen estará formado por los primeros 250 caracteres del cuerpo del texto.
- BC: realiza un índice a partir de los contenidos asignados a la propiedad «keywords» de la etiqueta META de HTML. Al mismo tiempo, crea diferentes ficheros invertidos con la información extraída de la etiqueta de título, los encabezamientos de nivel uno (H1), los campos de dirección <ADDRESS> y la URL.
- Otro que ya no existe como motor de búsqueda, Ultraseek, sufrió diferentes modificaciones en sus últimas implementaciones con el fin de permitir la indización y posterior recuperación a través de meta-etiquetas, ya sean las procedentes de HTML o las propias de otro modelo (especialmente Dublin Core).
- Aliweb: es un buscador que está basado en la indización distribuida por medio del uso del formato IAFA (Internet Anonymous FTP Archives).
- Harvest: desarrolló el sistema SOIF (Summary Object Interchange Format).
- En el mercado es posible encontrarse con programas que son capaces de indizar satisfactoriamente documentos HTML que incluyan metaetiquetas, como por ejemplo: Microsoft Site Index, Autonomy Knowledge Server, Blue Angel MetaStar, Inktomi o Verity K2 Enterprise.

Aunque hay que reconocer que existen buscadores e índices que no utilizan las etiquetas META en ningún momento del proceso de indización o búsqueda (caso de About, Ask Jeeves...), la mayoría de los servicios de recuperación de información en Internet están adaptando sus motores a sistemas de metadatos que ofrecen una información más completa del documento (como es el caso de Dublin Core o RDF).

2 Objetivos

Resulta evidente que es necesario utilizar todas las herramientas posibles para lograr superar la barrera impuesta por el enorme número de páginas web que impide lo-

calizar información de forma precisa. Estas herramientas y esfuerzos no deberían centrarse sólo en lograr un efectivo diseño corporativo o en cumplimentar el formulario de petición de alta en un buscador o índice determinado. Es necesario apoyar esta estrategia con otros elementos que permitan posicionar favorablemente un sitio web dentro de la lista de resultados de un servicio de consulta en Internet.

Teniendo en cuenta que el lenguaje utilizado para representar información en la web faculta la opción de realizar descripciones por medio un sistema de metadatos propio, podemos afirmar que la etiqueta META es, entre otras, una de las herramientas más apropiadas para lograr este objetivo.

Son pocos los estudios que se han publicado sobre el uso de la etiqueta META del HTML en documentos web. Los realizados por Lawrence (4) y publicados en Nature indican que sólo el 34% de las «homepages» cuentan con las etiquetas que contienen datos relativos a las palabras clave y al resumen (en este mismo estudio se afirma que el 0,3% de sitios web emplean las etiquetas Dublin Core). Craven (2), por su parte, realizó un estudio donde se plasma que algo menos del 26% de las páginas principales analizadas (en su mayoría pertenecientes a instituciones públicas) utilizan la etiqueta «description» (de este porcentaje se eliminan, obviamente, aquellas en que tienen mal utilizada dicha etiqueta). Otros autores (11) examinan el impacto de estos metadatos en empresas privadas. Seleccionando los sitios web pertenecientes a las 30 primeras y las 30 últimas empresas aparecidas en la lista Fortune se determinó que el 37% en el año 2000 y el 43% un año más tarde utilizan este tipo de herramientas para lograr una mayor visibilidad de las compañías dentro de Internet.

En el capítulo español destacan las aportaciones de Rodríguez Gairín (12) sobre la valoración del impacto de la información en Internet empleando motores de búsqueda, o los trabajos de Vidal Borges y Salvador Oliván (13) por un lado, y Merlo Vega y Sorli Rojo (14) por otro. Precisamente, este texto puede servir para, entre otras cosas, actualizar la información aportada por estos autores y comprobar los cambios acaecidos en este particular durante los últimos años.

En general, estos trabajos evidencian que la teoría de Craven y Turner al respecto de la importancia de las etiquetas META para alcanzar un mejor posicionamiento dentro de los resultados ofrecidos por los motores de búsqueda es realmente cierta. Si bien es cierto que el número de sitios web que cuenta con estas etiquetas es, según estos estudios, relativamente bajo (oscila entre el 26% y el 43%), no es menos cierto que la etiqueta META se contempla como un elemento muy importante para facilitar la localización de información (homepages institucionales, en este caso) dentro de Internet.

Partiendo de este hecho, y estando patente la idea de que los profesionales de la información no son ajenos a esta realidad, es lógico pensar que dentro de una profesión caracterizada —entre otras muchas cosas— por la concienciación sobre la normalización, la importancia de establecer mecanismos que faciliten la recuperación de la información y la habilidad demostrada para trabajar con normas internacionales relacionadas con la descripción de objetos (libros, revistas, páginas web...), no parece por tanto baladí plantear la hipótesis de que las páginas web de las bibliotecas deberían moverse, como mínimo, entre unos valores muy parecidos a los anteriormente comentados en lo que se refiere a la utilización de instrumentos que permitan localizar sus sitios web dentro de Internet.

Se pretende, por lo tanto, analizar si las páginas web de los centros bibliotecarios

disponen de los mecanismos adecuados que permitan describir su contenido, palabras clave, persona responsable...

3 Metodología

Para realizar este estudio se han constituido cuatro grupos de bibliotecas, atendiendo a su tipología, objetivos y nacionalidad. Dado que es prácticamente imposible analizar todos los sitios web de bibliotecas se optó por seleccionar una muestra suficientemente representativa. El primer grupo está formado por 59 bibliotecas universitarias españolas (tabla 1). El segundo grupo lo constituyen 60 sitios web de bibliotecas de estudios superiores (universidad y college) norteamericanas (tabla 2). Dado que en Estados Unidos existe una gran cantidad de instituciones de este tipo, se optó por escoger, de forma aleatoria, 10 centros de cada una de las seis particiones geográficas norteamericanas (Northeast, Midwest, Southeast, Southwest, Mountain and Plains States y West). El tercer conjunto lo componen las bibliotecas nacionales europeas más la americana Library of Congress, 38 en total (tabla 3). Por último, y con la finalidad de poder establecer una equiparación con el conjunto anterior, se encuentran 9 de las bibliotecas regionales españolas cabecera del sistema bibliotecario (tabla 4) (2). En total, el estudio se realizó sobre 166 bibliotecas de todo el mundo procurando que éstas fueran significativas dentro del entorno en el que desarrollaban su actividad.

Teniendo en cuenta investigaciones anteriores, se optó por visionar sólo la página principal. Este punto está cimentado por la actuación llevada a cabo en todos los trabajos comentados en el punto anterior así como las conclusiones alcanzadas en el estudio de Craven, según el cual, en la mayoría de sitios web se repite en todas las páginas la información almacenada en las metaetiquetas de la primera, salvo pequeñas modificaciones en aquellos documentos que cuentan con características muy divergentes con respecto al resto.

Tampoco se han contado como válidas aquellas etiquetas mal empleadas (caso que se ha dado en 6 ocasiones) o no cumplimentadas (en 3 ocasiones). La visualización de las páginas se ha realizado durante la tercera semana del mes de marzo de 2003, utilizando la opción «código fuente» de la versión 6 del navegador Internet Explorer.

4 Resultados

De las diferentes propiedades que pueden utilizarse con la etiqueta META, es «name» la que permite especificar con mayor claridad características propias de la descripción de objetos. Por ese motivo, centramos gran parte del estudio en analizar su uso dentro de los cuatro grupos.

4.1 Uso de «Author»

Tal y como se observa en el gráfico 1, la mayoría de bibliotecas estudiadas presentan unos valores muy uniformes. Resulta especialmente llamativo que ninguna biblioteca regional española cuente con información relativa a la persona responsable de

Tabla 1. Bibliotecas universitarias españolas

<i>nombre</i>	<i>name</i>	<i>description</i>	<i>keyword</i>	<i>automáticas</i>	<i>dublin core</i>	<i>rdf</i>
Universidad de Almería	No	No	No	Sí	No	No
Universidad de Cádiz	No	No	No	Sí	No	No
Universidad de Córdoba	No	No	No	Sí	No	No
Universidad de Granada	No	No	No	Sí	No	No
Universidad de Huelva	No	No	No	Sí	No	No
Universidad de Jaén	No	No	No	No	No	No
Universidad de Málaga	No	No	No	Sí	No	No
Universidad de Sevilla	No	No	No	Sí	No	No
Universidad Internacional de Andalucía	No	No	No	Sí	No	No
Universidad Pablo de Olavide (Sevilla)	No	Sí	Sí	No	No	No
Universidad de Zaragoza	No	No	No	Sí	No	No
Universidad de Oviedo	Sí	No	No	No	No	No
Universidad de Oviedo	Sí	No	No	No	No	No
Universidad de La Laguna	No	No	No	Sí	No	No
Universidad de Las Palmas de Gran Canaria	Sí	No	No	No	No	No
Universidad de Cantabria	No	No	No	No	Sí	No
Universidad de Castilla La Mancha	No	No	No	No	Sí	No
Universidad Católica de Avila	No	No	No	Sí	No	No
Universidad de Burgos	No	No	No	Sí	No	No
Universidad de León	No	Sí	Sí	Sí	No	No
Universidad de Salamanca	No	No	No	Sí	No	No
Universidad Pontificia de Salamanca	No	No	No	Sí	No	No
Universidad de Valladolid	No	No	No	Sí	No	No
Universitat Autònoma de Barcelona	Sí	Sí	Sí	Sí	Sí	No
Universitat de Barcelona	Sí	Sí	Sí	No	No	No
Universitat de Girona	No	No	No	Sí	No	No
Universitat de Lleida	No	No	No	Sí	No	No
Universitat de Vic	No	No	No	Sí	No	No
Universitat Oberta de Catalunya	Sí	Sí	Sí	No	No	No
Universitat Politècnica de Catalunya	Sí	No	Sí	No	No	No
Universitat Pompeu Fabra	No	No	No	Sí	No	No
Universitat Ramon Llull	Sí	Sí	Sí	No	No	No
Universitat Rovira i Virgili	No	No	No	Sí	No	No
Universidad de Extremadura	No	No	No	Sí	No	No
Universidade da Coruña	No	No	No	Sí	No	No
Universidade de Santiago de Compostela	No	No	No	Sí	No	No
Universidade de Vigo	No	No	No	Sí	No	No
Universidad de La Rioja	No	Sí	Sí	No	No	No
Universidad de Alcalá de Henares	No	No	No	No	No	No
Universidad Autónoma de Madrid	No	No	No	No	No	No
Universidad Carlos III	No	No	No	No	No	No
Universidad Complutense	No	No	No	No	Sí	No
Universidad Nacionbal de Educación a Distancia	No	No	No	No	No	No
Universidad Europea de Madrid	No	No	No	No	Sí	No
Universidad Politécnica de Madrid	Sí	Sí	Sí	No	No	No
Universidad Pontificia de Comillas	No	Sí	Sí	Sí	No	No
Universidad Rey Juan Carlos (Madrid)	No	No	No	Sí	No	No
Universidad Católica de San Antonio	No	No	No	Sí	No	No
Universidad de Murcia	Sí	No	No	Sí	No	No
Universidad Politécnica de Cartagena	No	No	No	No	No	No
Universidad de Navarra	No	No	No	Sí	No	No
Universidad Pública de Navarra	Sí	No	No	Sí	No	No
Universidad de Deusto	Sí	No	No	No	No	No
Universidad del País Vasco	No	Sí	No	No	No	No
Universitat d'Alacant	Sí	Sí	Sí	Sí	No	No
Universitat Cardenal Herrera	Sí	No	Sí	No	No	No
Universitat Jaume I	No	No	No	No	No	No
Universitat Politècnica de València	No	No	No	No	No	No
Universitat de València	No	No	No	Sí	No	No

Tabla 2. Bibliotecas de estudios superiores norteamericanos

<i>nombre</i>	<i>name</i>	<i>description</i>	<i>keyword</i>	<i>automáticas</i>	<i>dublin core</i>	<i>rdf</i>
University of Hartford	No	No	Sí	Sí	No	No
University of Massachusetts	No	No	No	Sí	No	No
University of New Hampshire	No	No	No	Sí	No	No
University of Pennsylvania	No	No	No	Sí	No	No
University of Pittsburgh	No	No	No	Sí	No	No
University of Rhode Island	No	Sí	Sí	Sí	No	No
University of Rochester	No	Sí	Sí	Sí	No	No
University of Vermont	No	No	No	Sí	No	No
Indiana School of Medicine	Sí	Sí	Sí	Sí	No	No
North Central University	No	No	No	Sí	No	No
Oakland City University	Sí	No	No	Sí	No	No
University of Chicago	Sí	No	No	Sí	No	No
University of Cincinnati	No	No	No	Sí	No	No
University of Evansville	No	No	No	Sí	No	No
University of Illinois	No	No	No	Sí	No	No
University of Iowa	No	No	No	Sí	No	No
American University	No	No	No	Sí	No	No
Bercia University	No	No	No	Sí	No	No
Columbia International University	No	Sí	Sí	No	No	No
Delaware State University	No	No	No	Sí	No	No
Georgetown University	Sí	Sí	Sí	Sí	No	No
North Carolina State University	No	No	No	No	No	No
University of Louisville	No	No	Sí	Sí	No	No
University of Virginia	No	No	No	Sí	Sí	No
Arizona State University	No	No	No	No	No	No
Houston Academy of Medicine	No	No	No	Sí	No	No
New Mexico State University	No	No	No	No	No	No
Palo Alto College	No	Sí	Sí	No	No	No
Rice University	No	No	No	Sí	No	No
University of Arizona	No	No	No	No	No	No
University of Dallas	No	No	No	No	No	No
University of Texas	No	No	No	No	No	No
Bellevue University	No	No	No	No	No	No
Dakota State University	No	No	No	Sí	No	No
Idaho State University	Sí	Sí	Sí	No	No	No
Montana State University	No	Sí	No	No	No	No
Preston University	No	Sí	Sí	No	No	No
University of Colorado	No	No	No	Sí	No	No
University of Denver	No	No	No	No	No	No
University of Nevada - Las Vegas	No	No	Sí	Sí	No	No
California Institute of Technology	No	No	No	Sí	No	No
California State University - Reno	Sí	No	No	Sí	No	No
Chapman University	No	No	No	Sí	No	No
Concordia University	No	No	No	No	No	No
University of California - Los Angeles	No	No	No	No	No	No
University of Oregon	No	No	No	Sí	No	No
University of Portland	No	No	No	Sí	No	No
University of the Pacific	No	No	No	Sí	No	No
Yakima Valley Community College	No	No	No	Sí	No	No
Yuba Community College	No	No	No	Sí	No	No
Williston State College	No	No	No	Sí	No	No
York College	No	No	No	Sí	No	No
Western Oklahoma State University	No	No	No	Sí	No	No
Yavapai College	Sí	No	No	Sí	No	No
York Technical College	No	No	No	Sí	No	No
Young Harris College	No	No	No	Sí	No	No
Wright State University	Sí	Sí	Sí	No	No	No
Youngstown State University	No	Sí	Sí	Sí	No	No
Yale University	No	No	No	No	No	No
York College	No	No	No	Sí	No	No

Tabla 3. Bibliotecas nacionales europeas más la americana Library of Congress

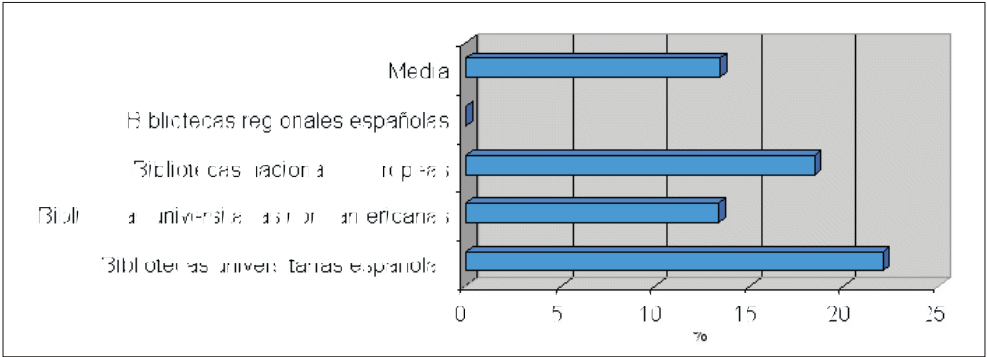
<i>nombre</i>	<i>name</i>	<i>description</i>	<i>keyword</i>	<i>automáticas</i>	<i>dublin core</i>	<i>rdf</i>
Austria	No	No	Sí	Sí	No	No
Bélgica	No	Sí	Sí	Sí	No	No
Bulgaria	No	No	No	No	No	No
Croacia	Sí	Sí	Sí	Sí	No	No
República Checa	No	No	No	Sí	Sí	No
Dinamarca	No	No	No	Sí	Sí	No
Estonia	No	No	Sí	Sí	No	No
Finlandia	No	No	No	Sí	No	No
Francia	No	Sí	Sí	Sí	Sí	No
Alemania	No	Sí	Sí	Sí	No	No
Grecia	No	No	No	Sí	No	No
Hungría	Sí	No	Sí	Sí	No	No
Islandia	No	No	No	Sí	Sí	No
Irlanda	No	No	No	No	No	No
Italia - Florencia	Sí	No	Sí	Sí	No	No
Italia - Roma	No	No	No	Sí	No	No
Liechtenstein	No	No	No	Sí	No	No
Lituania	No	Sí	Sí	Sí	No	No
Luxemburgo	No	No	No	Sí	No	No
Malta	Sí	Sí	Sí	Sí	No	No
Holanda	No	No	No	Sí	No	No
Noruega	Sí	Sí	Sí	Sí	Sí	No
Polonia	No	No	No	Sí	No	No
Portugal	No	No	No	Sí	No	No
Macedonia	No	No	No	Sí	No	No
Rumania	No	No	No	Sí	No	No
Rusia - Moscú	Sí	Sí	Sí	Sí	Sí	No
Rusia - San Petersburgo	No	Sí	Sí	No	No	No
Eslovaquia	No	No	No	No	No	No
Eslovenia	No	No	No	No	No	No
España	No	No	No	Sí	No	No
Suecia	No	Sí	No	Sí	Sí	No
Suiza	No	No	No	Sí	No	No
Turkía	Sí	Sí	Sí	Sí	No	No
Ucrania	No	Sí	Sí	Sí	No	No
Reino Unido	No	No	Sí	Sí	Sí	No
Ciudad del Vaticano	No	No	No	Sí	No	No
Bibliot. del Congreso (EE.UU.)	No	Sí	Sí	Sí	No	No

la puesta en marcha del servicio web o del mantenimiento del mismo. Algo que sí se manifiesta como una constante es que en la mayoría de sitios web esta etiqueta es utilizada para dar publicidad sobre la persona o, en la mayoría de ocasiones, empresa encargada de realizar las páginas web, y no así el webmaster. Por último, comentar que de las páginas que cuentan con esta etiqueta, apenas el 1% aprovecha este espacio para poner información relativa al correo electrónico u otros datos que faciliten la localización del responsable de la edición del documento, lo que da pie a pensar que en la mayoría de las ocasiones su utilización se limita a un hecho anecdótico y no a una política de visualización del personal por parte de la organización.

Tabla 4. Bibliotecas regionales españolas

<i>nombre</i>	<i>name</i>	<i>description</i>	<i>keyword</i>	<i>automáticas</i>	<i>dublin core</i>	<i>rdf</i>
Andalucía	No	No	No	No	No	No
Castilla La Mancha	No	No	No	Sí	No	No
Castilla y León	No	No	No	No	No	No
Cataluña	No	No	No	Sí	No	No
Galicia	No	No	No	Sí	No	No
Madrid	No	No	No	Sí	No	No
Murcia	No	No	No	Sí	No	No
Navarra	No	No	No	No	No	No
Valencia	No	No	No	No	No	No

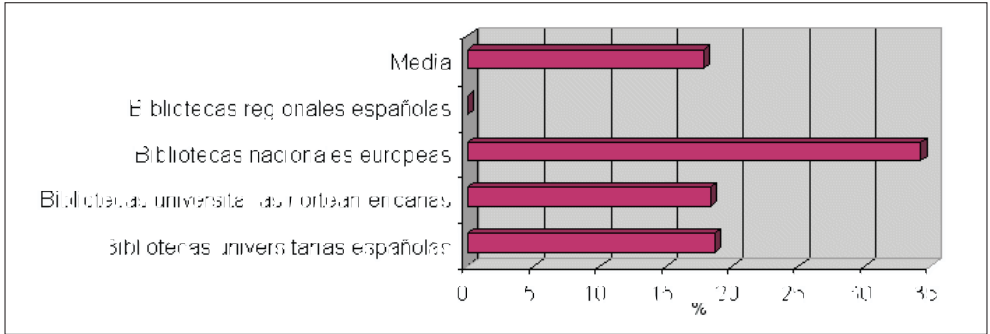
Gráfico 1. Uso de «author»



4.2 Uso de «Description»

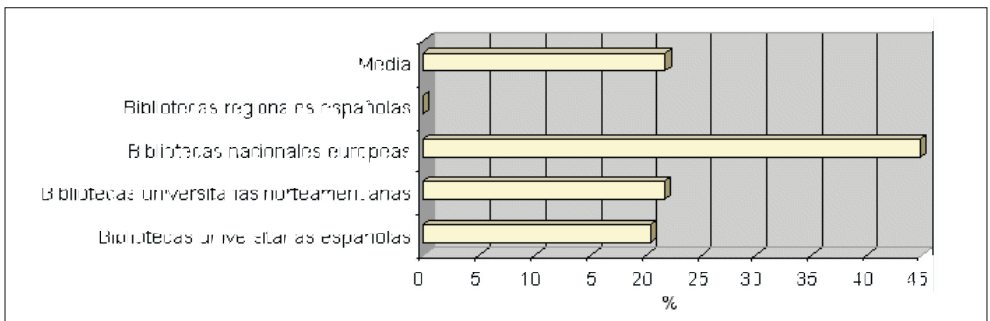
A pesar de que el uso correcto de esta etiqueta es fundamental para lograr una recuperación efectiva por parte de los motores de búsqueda y para poder gestionar mejor lo que aparece como resumen del sitio web dentro del listado de resultados de un buscador, no parece que las bibliotecas —salvo las englobadas dentro del grupo de nacionales europeas— le presten demasiada atención (gráfico 2). Las bibliotecas universitarias no se preocupan en exceso por realizar resúmenes en los que mostrar los diferentes servicios que ofrecen a sus usuarios. En muchos casos, la respuesta puede estar en que los sitios web de las bibliotecas universitarias se encuentran dentro de los sitios web de las universidades, y deberían ser éstas las que proporcionen una descripción más exhaustiva de sus servicios. En realidad esto no así, ya que implicaría contar con una política global a la hora de realizar el web corporativo que, obviamente, no se da. Además, es un hecho que se repite de forma sistemática independientemente desde donde se «cuelgue» la página web.

Mención aparte merecen, de nuevo, las bibliotecas regionales españolas, que tampoco emplean esta etiqueta demostrando una dejadez absoluta en lo que se refiere a la preocupación por ser visibles dentro de Internet. Por su parte, el grupo de las bibliotecas nacionales europeas se mantiene dentro de los porcentajes avanzados en estudios similares y explicados en el apartado 2 de este trabajo.

Gráfico 2. Uso de «description»

4.3 Uso de «Keywords»

Como media general, se puede observar un ligero avance con respecto al uso de las etiquetas comentadas hasta ahora (gráfico 3). En todos los casos se supera el 20% llegando, en el grupo de las nacionales europeas, a un esperanzador 44,7%. En este apartado no sólo se tiene que analizar la existencia de las etiquetas —lo que básicamente se mide en este estudio—, sino también la calidad y cantidad. Si bien es cierto que en algunas páginas nos hemos encontrado con no más de tres descriptores, en la mayoría de sitios, especialmente norteamericanos y europeos, la lista de palabras clave llegaba a contar con más de 20 términos, lo cual dice mucho sobre la preocupación de los responsables de estos sitios por ser fácilmente «recuperables» a través de los buscadores.

Gráfico 3. Uso de «Keywords»

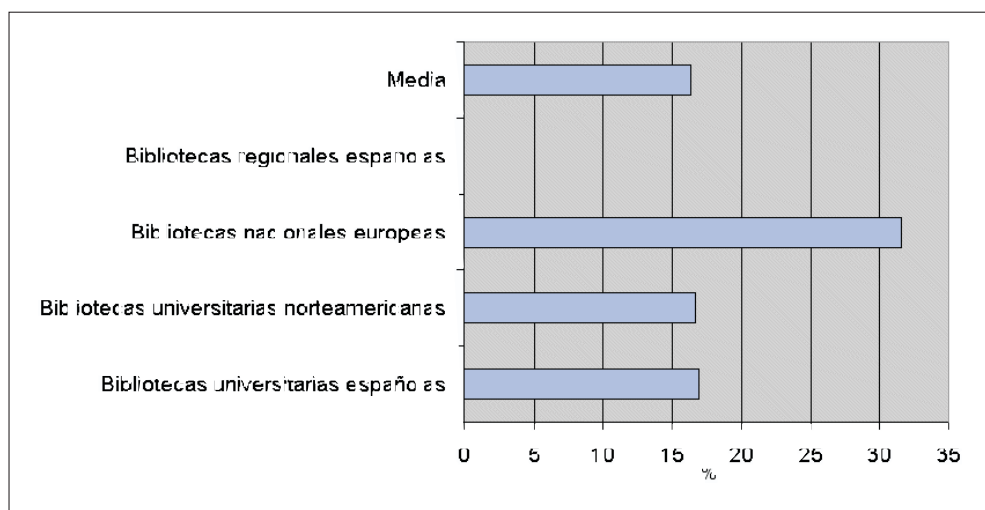
4.4 Combinaciones

De lo leído hasta ahora se deduce rápidamente que existen dos etiquetas que, juntas, aportan un nivel muy alto de descripción del documento en el que se insertan. «Description» y «Keywords» proporcionan los elementos necesarios para que cual-

quier página indizada por un buscador sea localizada de forma eficiente, evitando así el consabido ruido que generan las búsquedas en Internet y, además, proporciona cierto nivel de control sobre la información que aparecerá, al poder gestionar cada uno el contenido del resumen a mostrar. Tanto es así que la mayoría de estudios similares a este, realizados con diferentes conjuntos de páginas, hacen especial hincapié en la aparición de ambas etiquetas de forma conjunta.

En nuestro caso, gráfico 4, nos encontramos con que es muy complicado contar con un número razonable de centros que contemplen ambas posibilidades. Escasamente el 17% de las bibliotecas universitarias son capaces de hacer páginas web donde incluyan un resumen de los contenidos y, además, un conjunto significativo de palabras clave que facilite la recuperación en Internet. Por su parte, el 31% de los centros que constituyen el grupo formado por bibliotecas nacionales sí muestra un esfuerzo en esa línea. No obstante nos sigue pareciendo muy poco.

Gráfico 4. Uso de «description» y «keywords»



5 Conclusiones

El análisis de los datos ofrecidos tras el estudio de las páginas principales de los cuatro grupos de muestra de centros bibliotecarios (tabla 5) permite concluir que:

- La media de centros bibliotecarios que cuentan con etiquetas META del HTML lo suficientemente descriptivas del contenido ("description" y "keyword") en su página principal es del 16,3%, un valor muy por debajo de lo obtenido en estudios similares a nivel internacional (recordemos que se cifraba entre el 26% y el 43%). Aunque se supone que la nuestra es una profesión caracterizada, tal y como se comentó anteriormente, por la concienciación sobre la normalización, la importancia de establecer mecanismos que faciliten la recuperación de la información y la habilidad para trabajar con normas internacionales relacionadas con la descripción de objetos, es evidente que todos estos principios sólo se aplican del mostrador para atrás, e Internet se

Tabla 5. Resumen de los resultados

	author	descrip tion	key word	automá ticas	dc	description +keyword
Bibliotecas universitarias españolas	22%	18,6%	20,3%	57,6%	8,47%	16,9%
Bibliotecas universitarias norteamericanas	13,3%	18,3%	21,7%	71,7%	13,3%	16,7%
Bibliotecas nacionales europeas	18,4%	34,2%	44,7%	86,8%	2,63%	31,6%
Bibliotecas regionales españolas	0%	0%	0%	55,6%	0%	0%
Media	13,4%	17,8%	21,7%	67,9%	6,11%	16,3%

sigue viendo como una herramienta de consulta donde importa poco cómo sea la presencia que tenga la institución para la que se trabaja.

- Dentro de las bibliotecas universitarias españolas estudiadas existen sólo seis que cumplimentan las etiquetas "author", "description" y "keyword" (Universitat Autònoma de Barcelona, Universitat de Barcelona, Universitat Oberta de Catalunya, Universitat Ramon Llull, Universidad Politécnica de Madrid y Universitat d'Alacant). Curiosamente cuatro de ellas pertenecen a la comunidad de Cataluña. Creemos que este dato no se produce por azar sino que más bien es el reflejo de una realidad.

- Las bibliotecas universitarias estudiadas, ya sean nacionales o norteamericanas, presentan valores muy similares en todas las categorías, lo cual coloca a nuestros centros en un nivel equiparable al norteamericano, sin que eso signifique que sea un nivel aceptable.

- En general, el grupo formado por las bibliotecas nacionales europeas presenta unos niveles de uso de todas las etiquetas bastante aceptables y muy en consonancia con los datos mostrados en estudios similares con conjuntos de muestra diferentes (empresas, páginas personales, administración pública). Tanto es así que en casi todos los apartados la media global ha subido gracias a este grupo. Es especialmente reconfortante comprobar cómo algo más del 31% de centros incluye información relativa a palabras clave (en muchas ocasiones con listas muy extensas) y resumen de contenidos aunque, como ya hemos indicado, nos sigue pareciendo poco.

- Destaca la lamentable dejadez manifestada por todas las bibliotecas regionales estudiadas. Esto viene a demostrar, entre otras cosas, que en ninguno de estos centros existe una política corporativa de cara a Internet. En otras palabras, son firmes seguidores de una de las frases empresariales por excelencia de la década pasada: «hay que estar en Internet, aunque sólo sea por estar». Lo penoso es que esto sucedía hace 6 años y hoy las cosas han cambiado mucho aunque, lamentablemente, no para todos.

- Aunque este estudio no tiene como objetivo la exploración sobre la implantación de otros sistemas de metadatos dentro de los grupos de muestra seleccionados, se optó por aprovechar esta experiencia para avanzar algo sobre este tema. Así, descubrimos que el 8,4% de bibliotecas universitarias españolas, el 13,3% de bibliotecas universitarias norteamericanas y el 2,6% de las nacionales cuenta con una descripción de su sitio web basada en Dublin Core. Estos datos, que reflejan una media de uso total del 6,11%, ponen de manifiesto una realidad, y es que estos conjuntos de metadatos (que cada vez son más complejos de usar) tienen poco sentido en webs de este tipo. Su ubicación ideal se encuentra dentro de proyectos específicos como herramienta para facilitar la descripción más exhaustiva de objetos.

Por último, es necesario señalar el inconveniente que presenta el trabajar con este tipo de etiquetas que, en la mayoría de los casos, no están contempladas en los editores. Si bien es cierto que existe un numeroso grupo de programas que facilita la introducción de contenidos dentro de la etiqueta META (Meta Tag Carpenter, MetaLysar Lite, MetaMaster, Bhead y un largo etcétera) o plantillas en Internet, no hay tantos editores profesionales que faciliten la inclusión de este tipo de metadatos. Curiosamente casi el 68% de los sitios estudiados cuentan con etiquetas que hemos denominado «automáticas», es decir, generadas y anexadas al fichero directamente desde el propio editor web y que hacen mención al programa utilizado, el sistema de codificación empleado, etc. Sería interesante que las empresas propietarias de estas herramientas se plantearan la posibilidad de generar, también de forma automática, alguna de las etiquetas META.

6 Notas

1. Los servidores http pueden leer el contenido de la cabecera del documento y generar diferentes tipos de respuestas de acuerdo al valor definido en uno de los atributos que componen la etiqueta: HTTP-EQUIV.

2. No se han podido incluir más centros en este estudio ya que por motivos técnicos ha sido imposible el acceso a más centros.

7. Bibliografía

1. THELWALL, M. Commercial Web sites: lost in cyberspace? *Internet Research: Electronic Networking Applications and Policy*, 2000, vol. 10, n.º 2, pp. 150-159.
2. CRAVEN, T. C. Features of DESCRIPTION META tags in public home pages. *Journal of Information Science*, 2000, vol. 26, n.º 5, pp. 303-311.
3. CRAVEN, T. C. «Description» meta tags in locally linked web pages. *Aslib Proceedings*, junio 2001, vol. 53, n.º 6, pp. 203-216.
4. LAWRENCE, S. y GILE, L. Accessibility and Distribution of Information on the Web. *Nature*, 1999, vol. 400, pp. 107-109.
5. SULLIVAN, D. «Death of Meta Tag». *Search Engine Watch*, octubre 2002. <http://searchenginewatch.com/sereport/article.php/2165061>
6. GOOGMAN, A. «An End to Metatags (enough already)». *Traffik*, febrero 2002. <http://www.traffick.com/article.asp?aID=102>
7. TURNER, T. P. y BRACKBILL, L. Rising to the top: evaluating the use of the HTML meta tag to improve retrieval of WorldWideWeb documents through Internet search engines. *Library Resources and Technical Services*, 1998, vol. 42, n.º 4, pp. 258-271.
8. BERNERS-LEE, T. y CONNOLLY, D. *RFC 1866: Hypertext Markup Language - 2.0*, noviembre 1995. <http://www.ietf.org/rfc/rfc1866.txt?number=1866>
9. MENDEZ RODRÍGUEZ, E. *Metadatos y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea, 2002.
10. MARCKINI, F. *Search engine positioning*. Texas: Wordware Publishing, Inc., 2001.
11. DROTT, M. C. Indexing aids at corporate websites: the use of robots.txt and meta tags. *Information processing & management*, 2002, n.º 38, pp. 209-219.
12. RODRÍGUEZ GAIRÍN, J. M. Valoración del impacto de la información en Internet: Al-

vatista, el «citation index» de la Red. *Revista Española de Documentación Científica*, 1997, v. 20, n.º 2, pp. 175-181.

13. VIDAL BORDÉS, F. J.; SALVADOR OLIVÁN, J. A. La implementación de metadatos y Dublin Core en redes y páginas web de bibliotecas y centros de documentación de universidades y centros de investigación de la RedIRIS. *Jornadas Españolas de Documentación*. VII. 2000. Bilbao, 2000, pp. 197-209.
14. MERLO VEGA, J. A.; SORLI ROJO, A. El uso de metainformación en los webs de las bibliotecas españolas *Jornadas Españolas de Documentación*. VII. 2000. Bilbao, 2000, pp. 155-164.